DOCUMENT RESUME

ED 351 376 TM 019 202

AUTHOR Dowd, Steven B.
TITLE Multiple-Choice and Alternate-Choice Questions: Description and Analysis.
PUB DATE Sep 92
NOTE 16p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Comparative Testing; Difficulty Level; Evaluation Methods; Higher Education; *Multiple Choice Tests; Professional Education; *Radiologic Technologists; Radiology; Research Methodology; *Test Construction; Test Format; Test Items; Test Validity
IDENTIFIERS *Alternate Choice Questions

ABSTRACT

An alternative to multiple-choice (MC) testing is suggested as it pertains to the field of radiologic technology education. General principles for writing MC questions are given and contrasted with a new type of MC question, the alternate-choice (AC) question, in which the answer choices are embedded in the question in a short form that resembles real speech and the way in which material is taught. In a small study of the use of these questions, 10 students in a radiographic techniques course took a quiz in half MC and half AC formats. The quiz covered the initial lecture topic of radiographic density. Higher validity was found for the MC questions, which could be considered more difficult and better predictors of learning. Results do not support the value of AC questions, but it is suggested that radiologic technology educators can use a similar method to study the usefulness of AC questions. Three tables present the checklist for evaluating MC items and study findings. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED351376

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

[X] This document has been reproduced as received from the person or organization originating it.

[ ] Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

STEVEN B. DOWD

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Multiple-Choice and Alternate-Choice Questions:

Description and Analysis

Steven B. Dowd, Ed.D., R.T.(R)

TM019202

BEST COPY AVAILABLE

ABSTRACT

This article describes general principles for writing multiple-choice questions for radiologic technology education. Following this review, a description of a newer type of multiple-choice question, the alternate-choice question, is presented. A small study of the use of these question is described in a manner easily duplicated by radiologic technology instructors interested in the potential use of these questions.

INTRODUCTION

In radiologic technology, and the health professions in general, there is a strong focus on using multiple-choice test questions. This is due to a variety of factors, most probably the fact that registry and licensure boards make the most use of these types of questions. This article will review the writing of multiple-choice questions, some recent research on a "newer" form of multiple-choice question, the alternate-choice item, and review a small study of alternate choice questions in a radiography program that other instructors can duplicate to see if these types of questions have utility in their own programs.

WRITING MULTIPLE-CHOICE QUESTIONS

A multiple-choice question is a form of forced-choice (objective) question, consisting of a stem, which presents the problem, and a list of alternate answers including the correct response and distractors, which should be plausible, but incorrect or "not best."[1] The following rules must also be observed when writing multiple-choice items:

- Must consist of a problem.
- Alternate answers should be kept as short as possible, and of the same approximate length.
- Stems should consist of few negatives (after all, students tend to learn positive information; the do's rather than the don'ts more readily).
- Items should be independent of each other.
- Grammatical clues should be avoided.

- "None of the above" should not be used or only if it is either right or wrong. That is, it should not be used when writing a best answer question.
- Some authors also recommend against using of "all of the above" a student with only partial knowledge could get the item right. Yet sometimes clinical decisions must be made based on a partial, yet correct, knowledge of some of the facts.

A checklist that can be used when writing multiple-choice items can be found in Table I. Advantages and disadvantages of multiple-choice questions are listed in Table II.

The biggest problem from a professional standpoint of multiple-choice tests is their focus on simple recall of facts. Novice writers especially tend to write items that test recall.[2] A skilled instructor and item writer can write a good test using any test format - yet this can require considerable effort. Professional item writers for tests like the ACT are expected to write about ten items per day. Instructors often write a test within an hour or less. Such tests, even after years of testing, are prone to at least a few errors.

To evaluate student's clinical ability, standardized tests such as the registry use the test item known as K-type (complex multiple-choice item as opposed to the A-type, which consists of the single-best answer) to test for clinical ability. Yet research indicates that K-type items are not superior in

testing for complex cognitive knowledge[3] and thus, many licensing boards have discontinued their use. Ideally, some form of simulation or performance test to test for clinical knowledge would be used. The cost and manpower involved are obviously prohibitive.

THE ALTERNATE CHOICE QUESTION

In 1982 Ebel proposed a new type of test question he called **alternative choice** (AC).[4] Ebel noted that most clinical situations have as few as two choices, sometimes three. VanSustern, Cohen, and Simpson give the following as an example of an AC item:

> Ascending cholangitis usually occurs (a. before; b. after*) obstruction of the common bile duct.[5]

In radiography, when choosing technique factors, for example, the usual choices are between mAs and Kvp. A test question that asks about the use of distance to change density on a film might be theoretically correct but confusing to a student. A student must know that FFD changes will affect density but the clinical application of such a change is limited. For example:

> To increase density on a radiograph:
>
> 1. Increase Kvp
> 2. Increase mAs
> 3. Decrease distance (FFD)
>
> a. 1 and 2 only b. 1 and 3 only
>
> c. 2 and 3 only d. 1,2, and 3*

An example of an AC, best answer question similar to the

---

* - the asterisk indicates a correct response

3

above would be:

(a. Increasing kvp; b. Increasing mAs*) is the method usually preferred to increase optical density on a radiograph.

Similarly, this question:

The baseline between the external auditory meatus and the outer canthus of the eye is called the:

a. orbitomeatal line* b. acanthiomeatal line

c. mentomeatal line d. sagittal line

Would become, eliminating the less plausible distractors of mentomeatal line and sagittal line:

The (a. orbitomeatal line*; b. acanthiomeatal line) is the baseline between the external auditory meatus and the outer canthus of the eye.

The second choice is preferable to many instructors as it more closely resembles normal speech and how the information would be taught in the classroom.

Van Susteren tested the use of AC items with undergraduate psychology students and found that they were satisfactory in terms of difficulty, discrimination, and reliability as well as superior to A-type items in terms of testing for conceptual knowledge.[6] Maihoff and Mehrens took "old" multiple choice questions with four or five choices and eliminated all choices except for the correct answer and the most attractive distractor.[7] The items created were less difficult but not less discriminating in testing for knowledge

and application. This is advantageous in that the instructor can still determine the level of student understanding of the material while reducing the frustration level of students. Some instructors think that the more difficult a test is, quality of the test increases proportionally. Trickery, such as using double or even triple negatives, is counterproductive to learning.

The research test of AC questions closest to the discipline of radiologic technology has been VanSusteren, Cohen, and Simpson's use of the AC format with second-year medical students in a pathology course. They found that:

1. AC items were less difficult, but more reliable than. A- or K-type questions.
2. The AC format was more versatile in testing for clinical applications than other questions.
3. The free form style of the question allowed for a more gramatically natural question that reflected how the question would be asked during clinical practice. Thus, the question format was similar to how the material was taught.[5]

A critique of alternate choice questions could be that they are also,to a certain extent, true-false questions. However, clinical practice is also often a "yes-no," "more-less," higher-lower" decision-making process. They are also easy (and thus a relief) to write, as they are more easily extractable from lecture notes and texts, and resemble normal speech. Many

BEST COPY AVAILABLE

instructors are familiar with the dilemma of finding three plausible distractors. If alternate choice questions provide enough choices to determine student learning, then the process of writing questions would be facilitated.

A SMALL STUDY OF ALTERNATE CHOICE QUESTIONS

In 1976 Patricia Cross called for an increase in classroom research by instructors actually engaged in teaching to improve quality of instruction and reshape curriculum.[8] She noted that community college educators in particular have the resources to determine effective classroom learning strategies through applied research. Such research should be relevant to, and easily duplicated by, other instructors engaged in classroom teaching. This small study is designed to determine the applicability of AC questions in one setting and to present that research in a manner duplicable by other educators not only in this, but in other settings as well.

Ten students participating in a first-year, third semester course in Radiographic Technique formed the sample group. One of the regular quizzes for the course was specially designed as one-half regular multiple-choice (10 questions); and one-half alternate-choice (10 questions). This quiz covered the initial lecture topic of Radiographic Density. The quiz was analyzed for face validity by faculty members from another CAHEA-accredited radiography program.

The average score on the quiz was 14.4 (72.2%), with a standard deviation of 2.58 (Table III). This is consistent with

scores previously achieved by this group of students in courses taught by this instructor.

The test was also scored separately for the AC and multiple choice questions. The average score on the AC questions was 8.0 (80%), with a standard deviation of 1. The average score on the section of regular multiple-choice questions was 6.4 (64%), with a standard deviation of 2.32 .

The rank difference correlation (Spearman rho) was used to determine the relationship (r) between the two sets of scores. This test was calculated as described in Isaac and Michael (pg. 172).[9] A correlation coefficient is the statistical method of choice in this case as the relationship between variables is of primary importance. It measures to what extent variables in one factor (scores on the AC items) go with with variations in another (multiple-choice items). Thus scores on the AC items should correlate with multiple-choice items if both items measure student learning of a topic.

The value of r was 0.23, which was not statistically significant at either the .05 or .01 levels (8 df). A correlation coefficient ranges from -1.00, in which case changes in one variable are accompanied by equivalent changes in the opposite direction of the other variable, to 1.00. A perfect positive correlation of 1.00 indicates that changes in one variable are accompanied by equivalent changes in the same direction of the other variable. The weak positive correlation indicates that there is some positive association, but it could

well be attributed to chance.

This indicates that it cannot be assumed, with reasonable certainty, that both types of questions possess similar validity in measuring student learning in the population. The four-choice multiple-choice questions could be considered more difficult, as well as possibly a better predictor of learning. An item analysis using chi-square (Isaac and Michael, pp. 116-117) indicated that only one of the AC items should be retained, whereas eight of the multiple-choice items should be retained.

CONCLUSION

In this small study, AC questions were not found to be of value. This study was limited in its focus, and was not intended to answer broad-scale questions about AC items such as their applicability to clinical judgment or their use in other settings. It is certainly possible, for example, that the questions used in the AC section were simply comparatively easy for students. One planned follow-up to this study is to develop a test similar to the one developed by Maihoff and Mehrens, eliminating the two most implausible distractors.

Radiologic technology educators can use a similar method to the one described here of developing a test with equal numbers of two types of questions and calculating a correlation coefficient to determine a relationship between the two types. This can be performed with any two types of test items. Some educators feel, for example, that essay questions provide a good gauge of clinical knowledge. This could be evaluated in a

similar manner. Knowing what types of questions are the most effective will be of value to the educator wanting to evaluate all types of student learning.

References

1. Holcomb D. Test construction. In Ford CW and Morgan MK (Eds.): Teaching in the Health Professions, St. Louis, C.V. Mosby, 1976. pp. 198-207.

2. Ebel RL: Essentials of Educational Measurement. Englewood Cliffs, N.J.: Prentice-Hall, 1979.

3. Dawson-Saunders B, Nungester RJ, and Downing SM: A comparison of single best answer multiple-choice items (A-type) and complex multiple-choice items (K-type). Presented at the 28th Annual Conference of the Association of American Medical Colleges, 1989.

4. Ebel RL: Proposed solutions to two problems of test construction. Journal of Educational Measurement 1982; 19: 267-277.

5. Van Susteren TJ, Cohen EB, and Simpson DE: Alternate-choice test items: implications for measuring clinical judgment. Teaching and Learning in Medicine 1991; 3(1): 33-37.

6. Van Susteren TJ: The comparative reliabilities and validities of alternate choice and multiple choice items. Paper presented at the meeting of the American Educational Research Association, Boston, April 1990.

7. Maihoff NA and Mehrens WA: A comparison of multiple-choice and alternate-choice item forms on classroom tests. Paper presented at the meeting of the American Educational Research Association, Boston, April 1990.

8. Cross KP: Accent on Learning: Improving Instruction and Reshaping the Curriculum. San Francisco: Jossey-Bass, 1976.

9. Isaac S and Michael WB: Handbook in Research and Evaluation. San Diego: EDITS Publishers, 1981.

BEST COPY AVAILABLE

Table 1

Checklist for Evaluating Multiple-Choice Items*

Check the Suggestions you feel you have followed.

_____ When possible, the stem was stated as a direct question rather than as an incomplete statement.

_____ A definite, explicit, and singular question or problem was stated in the stem.

_____ Excessive verbiage and irrelevant information was eliminated from the stem.

_____ Included in the stem word(s) that would otherwise be repeated in each, or a few items.

_____ Negatively stated items used sparingly, and underlined when used.

_____ Made all items plausible.

_____ Alternatives are grammatically parallel and consistent with stem.

_____ When possible, alternatives are presented in logical order - most to least or some other order.

_____ Only one correct or best response (A-type item).

_____ Alternatives are approximately equal in length.

_____ Used "none of the above" or "all of the above" sparingly. If used, they are occasionally correct.

_____ Attempted to balance the number of correct responses between choices of a,b,c,d, etc.

* Adapted from Ory JC: Improving Your Test Questions. Urbana, IL: Office of Instructional and Management Services of the University of Illinois, 1983.

Table II

Advantages and Disadvantages of Multiple-Choice Questions

Advantages	*Disadvantages*
Versatile in measuring all levels of cognitive ability and content.	Difficult and time-consuming to construct.
Very reliable test scores.	Favor simple recall of facts.
Efficient and accurate to score.	Place high dependence on student reading ability.
Provide diagnostic feedback based on response to distractors.	Instructor must be a good writer.

Table III

Results

Average score (20 items)	14.4 (sd = 2.58)
Average score (10 AC items)	8.0 (sd = 1)
Average score (10 multiple-choice items)	6.4 (sd = 2.32)
r	.23*

* not significant at the .05 or .01 level